The Symmetric Solution of the Matrix Equations* $AX + YA = C$, $AXA^T + BYB^T = C$, and $(A^TXA, B^TXB) = (C, D)$

Xiao-Wen Chang and Jia-Song Wang
Department of Mathematics
Nanjing University
Nanjing 210008, P. R. China

Submitted by Richard A. Brualdi

ABSTRACT

The necessary and sufficient conditions for the existence of and the expressions for the symmetric solutions of matrix equations (I) $AX + YA = C$, (II) $AXA^T + BYB^T = C$, and (III) $(A^TXA, B^TXB) = (C, D)$ are derived. In addition, the minimum-2-norm least-squares symmetric solution of equation (I), the minimum-2-norm symmetric solution of equation (II), and the least-squares solution of equation (III) are obtained.

1. INTRODUCTION

Let $R^{m\times n}$ denote the class of real $m \times n$ matrices; $\mathrm{SR}^{n\times n}$, the class of real symmetric $n \times n$ matrices; $\mathrm{OR}^{n\times n}$, the class of real orthogonal $n \times n$ matrices. $\|\cdot\|_F$ stands for the Frobenius norm of a matrix. $A * B$ represents the Hadamard product of two $n \times n$ matrices A and B, that is, $A * B = (a_{ij}b_{ij})_{1\leqslant i,j\leqslant n}$.

The symmetric solutions of some matrix equations have been investigated by Vetter [10], Khatri and Mitra [4], Magnus and Neudecker [5–7], Don [2],

*Research supported by National Natural Science Foundation of China.

655 Avenue of the Americas, New York, NY 10010 0024-3795/93/$5.00

Chu [1], Wang and Chang [11], and others. In this paper, the following problems are considered.

PROBLEM I. Given $A \in R^{m\times n}$, $C \in R^{m\times n}$.

(a) Let

$$\mathscr{L}_{\mathrm{ILS}} = \{[X,Y] : X \in \mathrm{SR}^{n\times n},\ Y \in \mathrm{SR}^{m\times m},\ \|AX + YA - C\|_F = \min\}. \tag{1.1}$$

Find $[\hat{X},\hat{Y}] \in \mathscr{L}_{\mathrm{ILS}}$ such that

$$\|[\hat{X},\hat{Y}]\|_F \equiv \left(\|\hat{X}\|_F^2 + \|\hat{Y}\|_F^2\right)^{\frac{1}{2}} = \min. \tag{1.2}$$

(b) Let

$$\mathscr{L}_{\mathrm{I}} = \{[X,Y] : X \in \mathrm{SR}^{n\times n},\ Y \in \mathrm{SR}^{m\times m},\ AX + YA = C\}. \tag{1.3}$$

Find $[\hat{X},\hat{Y}] \in \mathscr{L}_{\mathrm{I}}$ such that

$$\|[\hat{X},\hat{Y}]\|_F = \min. \tag{1.4}$$

PROBLEM II. Given $A \in R^{m\times n}$, $B \in R^{m\times p}$, $C \in R^{m\times m}$. Let

$$\mathscr{L}_{\mathrm{II}} = \{[X,Y] : X \in \mathrm{SR}^{n\times n},\ Y \in \mathrm{SR}^{p\times p},\ AXA^T + BYB^T = C\}. \tag{1.5}$$

Find $[\hat{X},\hat{Y}] \in \mathscr{L}_{\mathrm{II}}$ such that

$$\|[\hat{X},\hat{Y}]\|_F = \min. \tag{1.6}$$

PROBLEM III. Given $A \in R^{m\times n}$, $B \in R^{m\times p}$, $C \in R^{n\times n}$, $D \in R^{p\times p}$.

(a) Find

$$\mathscr{L}_{\mathrm{III\,LS}} = \{X : X \in \mathrm{SR}^{m\times m},\ \|[A^TXA - C, B^TXB - D]\|_F = \min\}. \tag{1.7}$$

(b) Find

$$\mathscr{L}_{\mathrm{III}} = \{X : X \in \mathrm{SR}^{m \times m},\ A^T X A = C,\ B^T X B = D\}. \tag{1.8}$$

Using singular-value and generalized singular-value decompositions, the necessary and sufficient conditions for the existence of solutions of Problems I(b), II, and III(b) are obtained, and expressions for the solutions of Problems I, II, and III are also given.

For convenience of discussion in the later sections, here we give the singular-value decomposition (SVD) of a matrix A, and the generalized singular value decomposition (GSVD) of a matrix pair $[A, B]$.

Given a matrix $A \in R^{m \times n}$ of rank r, its SVD is

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T, \tag{1.9}$$

where

$$U = \Big(\underset{r}{U_1},\ \underset{m-r}{U_2} \Big) \in \mathrm{OR}^{m \times m}, \qquad V = \Big(\underset{r}{V_1},\ \underset{n-r}{V_2} \Big) \in \mathrm{OR}^{n \times n},$$

$$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) > 0.$$

Given two matrices $A \in R^{m \times n}$, $B \in R^{m \times p}$, the GSVD of $[A, B]$ is

$$A = M \Sigma_A U^T, \qquad B = M \Sigma_B V^T, \tag{1.10}$$

where M is a nonsingular $m \times m$ matrix,

$$U = \Big(\underset{r}{U_1},\ \underset{s}{U_2},\ \underset{n-r-s}{U_3} \Big) \in \mathrm{OR}^{n \times n}, \qquad V = \Big(\underset{p+r-k}{V_1},\ \underset{s}{V_2},\ \underset{k-r-s}{V_3} \Big) \in \mathrm{OR}^{p \times p},$$

and

$$\Sigma_A = \begin{array}{c} \begin{pmatrix} I_A & & \\ & S_A & \\ & & O_A \\ \hdashline & 0 & \end{pmatrix} \end{array} \begin{array}{l} r \\ s \\ k-r-s \\ m-k \end{array}, \tag{1.11}$$

$$\Sigma_B = \begin{array}{c} \begin{pmatrix} O_B & & \\ & S_B & \\ & & I_B \\ \hdashline & 0 & \end{pmatrix} \end{array} \begin{array}{l} r \\ s \\ k-r-s \\ m-k \end{array}. \tag{1.12}$$

Here, $k = \operatorname{rank} C = \operatorname{rank}(A, B)$, $r = k - \operatorname{rank} B$, $s = \operatorname{rank} A + \operatorname{rank} B - k$, I_A and I_B are identity matrices, O_A and O_B are zero matrices, and $S_A = \operatorname{diag}(\alpha_1, \dots, \alpha_s)$, $S_B = \operatorname{diag}(\beta_1, \dots, \beta_s)$ with $1 > \alpha_1 \geqslant \cdots \geqslant \alpha_s > 0$, $0 < \beta_1 \leqslant \cdots \leqslant \beta_s < 1$, $\alpha_i^2 + \beta_i^2 = 1$, $i = 1, \dots, s$.

Some submatrices in (1.11) and (1.12) may disappear, depending on the structure of the matrices A and B.

Proofs and algorithms about the GSVD can be found in [3, 8, 9].

2. SOLVING PROBLEM I

Let us first introduce a lemma.

LEMMA 2.1. *Given $G \in R^{r\times r}$, $\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_r) > 0$. Define*

$$\varphi_{ij} = \begin{cases} \dfrac{1}{\sigma_i^2 - \sigma_j^2} & \sigma_i \neq \sigma_j, \\ 0 & \sigma_i = \sigma_j, \end{cases} \quad 1 \leqslant i, j \leqslant r, \qquad \Phi = (\varphi_{ij}) \in R^{r\times r}, \tag{2.1}$$

$$\psi_{ij} = 1 - \left|\operatorname{sgn}(\sigma_i - \sigma_j)\right|, \quad 1 \leqslant i, j \leqslant r, \qquad \Psi = (\psi_{ij}) \in R^{r\times r}. \tag{2.2}$$

Let

$$\mathscr{L}_{\mathrm{LS}} = \left\{[S, T] : S \in SR^{r\times r}, T \in SR^{r\times r}, \|\Sigma S + T\Sigma - G\|_F = \min\right\}, \tag{2.3}$$

$$\mathscr{L} = \left\{[S, T] : S \in SR^{r\times r}, T \in SR^{r\times r}, \Sigma S + T\Sigma = G\right\}. \tag{2.4}$$

Then

(1) *The expression for $\mathscr{L}_{\mathrm{LS}}$ is*

$$\mathscr{L}_{\mathrm{LS}} = \Big\{\big[\Phi * (\Sigma G - G^T\Sigma) + \Psi * \big(\tfrac{1}{2}\Sigma^{-1}(G + G^T) - T\big), \\ \Phi * (\Sigma G^T - G\Sigma) + \Psi * T\big], T \in \mathrm{SR}^{r\times r}\Big\}. \tag{2.5}$$

In set $\mathscr{L}_{\mathrm{LS}}$, there exists a unique $[\hat{S}, \hat{T}]$ such that

$$\left\|[\hat{S}, \hat{T}]\right\|_F = \min, \tag{2.6}$$

and $[\hat{S}, \hat{T}]$ *can be expressed as*

$$\left[\hat{S}, \hat{T}\right] = \left[\Phi * (\Sigma G - G^T\Sigma) + \tfrac{1}{4}\Psi * \left(\Sigma^{-1}(G + G^T)\right),\right.$$
$$\left.\Phi * (\Sigma G^T - G\Sigma) + \tfrac{1}{4}\Psi * \left(\Sigma^{-1}(G + G^T)\right)\right]. \quad (2.7)$$

(2) *The set* $\mathscr{L}$ *is nonempty if and only if*

$$\Psi * (G - G^T) = 0. \quad (2.8)$$

Under this condition, $\mathscr{L}$ *can be expressed as*

$$\mathscr{L} = \left\{\left[\Phi * (\Sigma G - G^T\Sigma) + \Psi * (\Sigma^{-1}G - T), \Phi * (\Sigma G^T - G\Sigma) + \Psi * T\right],\right.$$
$$\left. T \in \mathrm{SR}^{r\times r}\right\}. \quad (2.9)$$

There exists a unique $[\hat{S}, \hat{T}]$ *in* $\mathscr{L}$ *such that*

$$\left\|\left[\hat{S}, \hat{T}\right]\right\|_F = \min, \quad (2.10)$$

and $[\hat{S}, \hat{T}]$ *can be expressed as*

$$\left[\hat{S}, \hat{T}\right] = \left[\Phi * (\Sigma G - G^T\Sigma) + \tfrac{1}{2}\Psi * (\Sigma^{-1}G),\right.$$
$$\left.\Phi * (\Sigma G^T - G\Sigma) + \tfrac{1}{2}\Psi * (\Sigma^{-1}G)\right]. \quad (2.11)$$

Proof. (1): For $S = (s_{ij}) \in \mathrm{SR}^{r\times r}$, $T = (t_{ij}) \in \mathrm{SR}^{r\times r}$, and $G = (g_{ij}) \in R^{r\times r}$, we have

$$\|\Sigma S + T\Sigma - G\|_F^2 = \sum_{1\leqslant i\leqslant r} (\sigma_i s_{ii} + \sigma_i t_{ii} - g_{ii})^2$$
$$+ \sum_{1\leqslant i<j\leqslant r} \left[\left(\sigma_i^2 + \sigma_j^2\right)s_{ij}^2 + \left(\sigma_i^2 + \sigma_j^2\right)t_{ij}^2\right.$$
$$-2(\sigma_i g_{ij} + \sigma_j g_{ji})s_{ij} - 2(\sigma_i g_{ji} + \sigma_j g_{ij})t_{ij}$$
$$\left.+4\sigma_i\sigma_j s_{ij}t_{ij} + g_{ij}^2 + g_{ji}^2\right]. \quad (2.12)$$

From (2.12), it is easy to deduce that $[S, T] \in \mathscr{L}_{\mathrm{LS}}$ should satisfy

$$s_{ij} = \frac{\sigma_i g_{ij} - \sigma_j g_{ji}}{\sigma_i^2 - \sigma_j^2}, \; t_{ij} = \frac{\sigma_i g_{ji} - \sigma_j g_{ij}}{\sigma_i^2 - \sigma_j^2}, \quad \sigma_i \neq \sigma_j, \qquad 1 \leqslant i, j \leqslant r, \tag{2.13}$$

$$s_{ij} + t_{ij} = \frac{1}{2\sigma_i}(g_{ij} + g_{ji}), \quad \sigma_i = \sigma_j, \qquad 1 \leqslant 1, j \leqslant r. \tag{2.14}$$

Then (2.5) is obtained by (2.13) and (2.14).

For $[S, T] \in \mathscr{L}_{\mathrm{LS}}$,

$$\begin{aligned} \|[S, T]\|_F^2 &= \sum_{1 \leqslant i, j \leqslant r} \left(s_{ij}^2 + t_{ij}^2\right) \\ &= \sum_{\sigma_i \neq \sigma_j} \left(s_{ij}^2 + t_{ij}^2\right) + \sum_{\sigma_i = \sigma_j} \left(s_{ij}^2 + t_{ij}^2\right). \end{aligned} \tag{2.15}$$

On the right side of (2.15), the first term is definite. In order to minimize the left side of (2.15), by (2.14) s_{ij}, t_{ij} should satisfy

$$s_{ij} = t_{ij} = \frac{1}{4\sigma_i}(g_{ij} + g_{ji}), \quad \sigma_i = \sigma_j, \qquad 1 \leqslant i, j \leqslant r.$$

Thus $[\hat{S}, \hat{T}]$ is uniquely defined and given by (2.7).

(2): $\mathscr{L}$ is nonempty if and only if there exist $S \in \mathrm{SR}^{r \times r}$, $T \in \mathrm{SR}^{r \times r}$ such that

$$\left.\begin{aligned} \sigma_i s_{ij} + \sigma_j t_{ij} &= g_{ij} \\ \sigma_j s_{ij} + \sigma_i t_{ij} &= g_{ji} \end{aligned}\right\} \qquad 1 \leqslant i, j \leqslant r. \tag{2.16}$$

From (2.16), we know the necessary and sufficient condition under which $\mathscr{L}$ is nonempty is

$$g_{ij} = g_{ji}, \quad \sigma_i = \sigma_j, \qquad 1 \leqslant i, j \leqslant r, \tag{2.17}$$

and when that condition holds,

$$s_{ij} = \frac{\sigma_i g_{ij} - \sigma_j g_{ji}}{\sigma_i^2 - \sigma_j^2}, \; t_{ij} = \frac{\sigma_i g_{ji} - \sigma_j g_{ij}}{\sigma_i^2 - \sigma_j^2}, \quad \sigma_i \neq \sigma_j, \qquad 1 \leqslant i, j \leqslant r, \tag{2.18}$$

$$s_{ij} + t_{ij} = \frac{1}{\sigma_i} g_{ij}, \quad \sigma_i = \sigma_j, \qquad 1 \leqslant i, j \leqslant r. \tag{2.19}$$

Thus, (2.8) and (2.9) are proved.

Through a similar argument to that in the proof of (1), it is easy to verify that the minimum-2-norm solution $[\hat{X}, \hat{Y}]$ is uniquely defined by (2.11). ■

Now we give the following theorem.

THEOREM 2.1. *Let the singular-value decomposition of A given in Problem I be of the form* (1.9), *and let* Φ *and* Ψ *be defined by* (2.1) *and* (2.2) *respectively. Then* :

(*a*) *The set* $\mathscr{L}_{\mathrm{ILS}}$ *can be expressed as*

$$\mathscr{L}_{\mathrm{ILS}} = \left\{ \left[V \begin{pmatrix} \Phi * \left(\Sigma U_1^T C V_1 - V_1^T C^T U_1 \Sigma\right) & \\ \quad + \Psi * \left(\frac{1}{2}\Sigma^{-1}\left(U_1^T C V_1 + V_1^T C^T U_1\right) - Y_{11}\right) & \Sigma^{-1} U_1^T C V_2 \\ V_2^T C^T U_1 \Sigma^{-1} & X_{22} \end{pmatrix} V^T, \right.\right.$$

$$\left. U \begin{pmatrix} \Phi * \left(\Sigma V_1^T C^T U_1 - U_1^T C V_1 \Sigma\right) + \Psi * Y_{11} & \Sigma^{-1} V_1^T C^T U_2 \\ U_2^T C V_1 \Sigma^{-1} & Y_{22} \end{pmatrix} U^T \right],$$

$$\left. X_{22} \in \mathrm{SR}^{(n-r)\times(n-r)},\ Y_{11} \in \mathrm{SR}^{r\times r},\ Y_{22} \in \mathrm{SR}^{(m-r)\times(m-r)} \right\}. \tag{2.20}$$

In $\mathscr{L}_{\mathrm{ILS}}$, *there exists a unique* $[\hat{X}, \hat{Y}]$ that makes (1.2) hold and

$$\hat{X} = V \begin{pmatrix} \Phi * \left(\Sigma U_1^T C V_1 - V_1^T C^T U_1 \Sigma\right) & \Sigma^{-1} U_1^T C V_2 \\ + \frac{1}{4}\Psi * \left(\Sigma^{-1}\left(U_1^T C V_1 + V_1^T C^T U_1\right)\right) & \\ V_2^T C^T U_1 \Sigma^{-1} & 0 \end{pmatrix} V^T, \tag{2.21}$$

$$\hat{Y} = U \begin{pmatrix} \Phi * \left(\Sigma V_1^T C^T U_1 - U_1^T C V_1 \Sigma\right) & \Sigma^{-1} V_1^T C^T U_2 \\ + \frac{1}{4}\Psi * \left(\Sigma^{-1}\left(U_1^T C V_1 + V_1^T C^T U_1\right)\right) & \\ U_2^T C V_1 \Sigma^{-1} & 0 \end{pmatrix} U^T. \tag{2.22}$$

(b) *The set* $\mathscr{L}_I$ *is nonempty if and only if*

$$U_2^T C V_2 = 0, \qquad \Psi * \left(U_1^T C V_1 - V_1^T C^T U_1\right) = 0. \tag{2.23}$$

When the condition is satisfied, $\mathscr{L}_I$ has the following form:

$$\mathscr{L}_{\mathrm{I}} = \left\{ \left[V \begin{pmatrix} \Phi * \left(\Sigma U_1^T C V_1 - V_1^T C^T U_1 \Sigma\right) & \\ \quad + \Psi * \left(\Sigma^{-1} U_1^T C V_1 - Y_{11}\right) & \Sigma^{-1} U_1^T C V_2 \\ V_2^T C^T U_1 \Sigma^{-1} & X_{22} \end{pmatrix} V^T, \right. \right.$$

$$\left. U \begin{pmatrix} \Phi * \left(\Sigma V_1^T C^T U_1 - U_1^T C V_1 \Sigma\right) + \Psi * Y_{11} & \Sigma^{-1} V_1^T C^T U_2 \\ U_2^T C V_1 \Sigma^{-1} & Y_{22} \end{pmatrix} U^T \right],$$

$$\left. X_{22} \in \mathrm{SR}^{(n-r)\times(n-r)},\ Y_{11} \in \mathrm{SR}^{r\times r},\ Y_{22} \in \mathrm{SR}^{(m-r)\times(m-r)} \right\}. \tag{2.24}$$

In $\mathscr{L}_I$, there exists a unique $[\hat{X}, \hat{Y}]$ that makes (1.4) *hold.* $\hat{X}, \hat{Y}$ *can be expressed as*

$$\hat{X} = V \begin{pmatrix} \Phi * \left(\Sigma U_1^T C V_1 - V_1^T C^T U_1 \Sigma\right) + \frac{1}{2}\Psi * \left(\Sigma^{-1} U_1^T C V_1\right) & \Sigma^{-1} U_1^T C V_2 \\ V_2^T C^T U_1 \Sigma^{-1} & 0 \end{pmatrix} V^T, \tag{2.25}$$

$$\hat{Y} = U \begin{pmatrix} \Phi * \left(\Sigma V_1^T C^T U_1 - U_1^T C V_1 \Sigma\right) + \frac{1}{2}\Psi * \left(\Sigma^{-1} U_1^T C V_1\right) & \Sigma^{-1} V_1^T C^T U_2 \\ U_2^T C V \Sigma^{-1} & 0 \end{pmatrix} U^T. \tag{2.26}$$

Proof. (a): Using the SVD (1.9) of the matrix A, we have

$$\begin{aligned} \| AX + YA - C\|_F^2 &= \left\| U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T X + Y U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T - C \right\|_F^2 \\ &= \left\| \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T X V + U^T Y U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} - U^T C V \right\|_F^2. \end{aligned} \tag{2.27}$$

Write

$$V^T X V = \begin{array}{cc} \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} r & n-r \end{array} & \end{array}, \qquad U^T Y U = \begin{array}{cc} \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & Y_{22} \end{pmatrix} & \begin{array}{l} r \\ m-r \end{array} \\ \begin{array}{cc} r & m-r \end{array} & \end{array},$$

$$U^T C V = \begin{array}{cc} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} & \begin{array}{l} r \\ m-r \end{array} \\ \begin{array}{cc} r & n-r \end{array} & \end{array}. \tag{2.28}$$

Substitute (2.28) into (2.27):

$$\|AX + YA - C\|_F^2 = \|\Sigma X_{11} + Y_{11}\Sigma - C_{11}\|_F^2 + \|\Sigma X_{12} - C_{12}\|_F^2$$
$$+ \|Y_{12}^T \Sigma - C_{21}\|_F^2 + \|C_{22}\|_F^2. \tag{2.29}$$

Hence, for any $[X, Y] \in \mathscr{L}_{\mathrm{ILS}}$, the submatrices $X_{11} \in \mathrm{SR}^{r\times r}$, $X_{12} \in R^{r\times(n-r)}$, $Y_{11} \in \mathrm{SR}^{r\times r}$, and $Y_{12} \in R^{r\times(m-r)}$ should satisfy the following:

$$\|\Sigma X_{11} + Y_{11}\Sigma - C_{11}\|_F = \min, \tag{2.30}$$

$$\|\Sigma X_{12} - C_{12}\|_F = \min, \qquad \|Y_{12}^T\Sigma - C_{21}\|_F = \min. \tag{2.31}$$

Applying Lemma 2.1 to (2.30), we get

$$X_{11} = \Phi * \left(\Sigma C_{11} - C_{11}^T\Sigma\right) + \Psi * \left(\tfrac{1}{2}\Sigma^{-1}\left(C_{11} + C_{11}^T\right) - Y_{11}\right), \tag{2.32}$$

$$Y_{11} = \Phi * \left(\Sigma C_{11}^T - C_{11}\Sigma\right) + \Psi * Y_{11}. \tag{2.33}$$

From (2.31), we have

$$X_{12} = \Sigma^{-1} C_{12}, \qquad Y_{12} = \Sigma^{-1} C_{21}^T. \tag{2.34}$$

Therefore (2.20) is obtained from (2.28), (2.32), (2.33), and (2.34).

Obviously, $\mathscr{L}_{\rm ILS}$ is a closed convex set, so there exists a unique minimum-2-norm least-squares solution $[\hat{X}, \hat{Y}]$ in $\mathscr{L}_{\rm ILS}$. By using the expression for $\mathscr{L}_{\rm ILS}$ and Lemma 2.1, (2.21) and (2.22) are easily obtained.

(b): By (2.29) and Lemma 2.1, we can deduce that $\mathscr{L}_{\rm I}$ is nonempty if and only if

$$C_{22} = 0, \qquad \Psi * \left(C_{11} - C_{11}^T\right) = 0. \tag{2.35}$$

The above condition is equivalent to (2.23).

When the condition (2.35) is satisfied, (2.24), (2.25), and (2.26) are naturally obtained.

The theorem is proved. ∎

3. SOLVING PROBLEM II

First, we give a lemma, which will be used in this section and the next section.

LEMMA 3.1. *Given* $G \in R^{r\times r}$, $H \in R^{r\times r}$, $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_r) > 0$, $\Gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_r) > 0$, *there exists a unique matrix* $\hat{S} \in \mathrm{SR}^{r\times r}$ *such that*

$$\|\Lambda\hat{S}\Lambda - G\|_F^2 + \|\Gamma\hat{S}\Gamma - H\|_F^2 = \min_{S\in \mathrm{SR}^{r\times r}} \left(\|\Lambda S\Lambda - G\|_F^2 + \|\Gamma\hat{S}\Gamma - H\|_F^2\right) \tag{3.1}$$

and $\hat{S}$ *can be expressed as*

$$\hat{S} = \tfrac{1}{2}\Phi * \left(\Lambda(G + G^T)\Lambda + \Gamma(H + H^T)\Gamma\right), \tag{3.2}$$

where

$$\Phi = (\varphi_{ij}) \in R^{r\times r}, \qquad \varphi_{ij} = \frac{1}{\lambda_i^2\lambda_j^2 + \gamma_i^2\gamma_j^2}. \tag{3.3}$$

Proof. For $S = (s_{ij}) \in \mathrm{SR}^{r\times r}$, $G = (g_{ij}) \in R^{r\times r}$, and $H = (h_{ij}) \in R^{r\times r}$,

$$\begin{aligned}
&\|\Lambda S\Lambda - G\|_F^2 + \|\Gamma S\Gamma - H\|_F^2 \\
&\quad = \sum_{1\leqslant i\leqslant r} \left[(\lambda_i^4 + \gamma_i^4) s_{ii}^2 - 2(\lambda_i^2 g_{ii} + \gamma_i^2 h_{ii}) s_{ii} + g_{ii}^2 + h_{ii}^2 \right] \\
&\quad\quad + \sum_{1\leqslant i<j\leqslant r} \Big\{ 2(\lambda_i^2\lambda_j^2 + \gamma_i^2\gamma_j^2) s_{ij}^2 \\
&\quad\quad\quad - 2\left[\lambda_i\lambda_j(g_{ij} + g_{ji}) + \gamma_i\gamma_j(h_{ij} + h_{ji}) \right] s_{ij} \\
&\quad\quad\quad + g_{ij}^2 + g_{ji}^2 + h_{ij}^2 + h_{ji}^2 \Big\}.
\end{aligned} \tag{3.4}$$

From (3.4), it is easy to obtain a unique solution $\hat{S} = (\hat{s}_{ij}) \in \mathrm{SR}^{r\times r}$ of (3.1). $\hat{s}_{ij}$ is defined as

$$\hat{s}_{ij} = \frac{1}{2} \cdot \frac{\lambda_i(g_{ij} + g_{ji})\lambda_j + \gamma_i(h_{ij} + h_{ji})\gamma_j}{\lambda_i^2\lambda_j^2 + \gamma_i^2\gamma_j^2}, \qquad 1 \leqslant i, j \leqslant r. \tag{3.5}$$

Thus (3.2) is proved. ■

THEOREM 3.1. *Let the generalized singular-value decomposition of the matrix pair* $[A, B]$ *given in Problem* II *be of the form* (1.10). *Partition* $M^{-1}CM^{-T}$ *into the following form:*

$$M^{-1}CM^{-T} = \begin{array}{c} \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{pmatrix} \begin{array}{l} r \\ s \\ k-r-s \\ m-k \end{array} \\ \begin{array}{cccc} r & s & k-r-s & m-k \end{array} \end{array} \tag{3.6}$$

Then the set $\mathscr{L}_{\mathrm{II}}$ *is nonempty if and only if*

$$C = C^T, \qquad C_{13} = C_{14} = C_{24} = C_{34} = C_{44} = 0. \tag{3.7}$$

When the condition (3.7) *is satisfied,* $\mathscr{L}_{\mathrm{II}}$ *can be expressed as*

$$\mathscr{L}_{II} = \left\{ \left[U \begin{pmatrix} C_{11} & C_{12} S_A^{-1} & X_{13} \\ S_A^{-1} C_{12}^T & S_A^{-1}(C_{22} - S_B Y_{22} S_B) S_A^{-1} & X_{23} \\ X_{13}^T & X_{23}^T & X_{33} \end{pmatrix} U^T, \right. \right.$$

$$\left. V \begin{pmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{12}^T & Y_{22} & S_B^{-1} C_{23} \\ Y_{13}^T & C_{23}^T S_B^{-1} & C_{33} \end{pmatrix} V^T \right],$$

$$X_{13} \in R^{r \times (n-r-s)},\ X_{23} \in R^{s \times (n-r-s)},\ X_{33} \in \mathrm{SR}^{(n-r-s)\times(n-r-s)},$$

$$Y_{11} \in \mathrm{SR}^{(p+r-k)\times(p+r-k)},\ Y_{12} \in R^{(p+r-k)\times s},$$

$$\left. Y_{13} \in R^{(p+r-k)\times(k-r-s)},\ Y_{22} \in \mathrm{SR}^{s\times s} \right\}. \tag{3.8}$$

In $\mathscr{L}_{\mathrm{II}}$, *there exists a unique* $[\hat{X}, \hat{Y}]$ *that makes* (1.6) *hold, and* $\hat{X}, \hat{Y}$ *can be expressed as*

$$\hat{X} = U \begin{pmatrix} C_{11} & C_{12}^{-1} S_A & 0 \\ S_A^{-1} C_{12}^T & \Phi * (S_A C_{22} S_A) & 0 \\ 0 & 0 & 0 \end{pmatrix} U^T,$$

$$\hat{Y} = V \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Phi * (S_B C_{22} S_B) & S_B^{-1} C_{23} \\ 0 & C_{23}^T S_B^{-1} & C_{33} \end{pmatrix} V^T, \tag{3.9}$$

where

$$\Phi = (\varphi_{ij}) \in R^{s\times s}, \qquad \varphi_{ij} = \frac{1}{\alpha_i^2 \alpha_j^2 + \beta_i^2 \beta_j^2}. \tag{3.10}$$

Proof. If the set $\mathscr{L}_{\mathrm{II}}$ is nonempty, obviously, C must be symmetric. For $[X, Y] \in \mathscr{L}_{\mathrm{II}}$, using the GSVD (1.10) of $[A, B]$, we have

$$M\Sigma_A U^T X U \Sigma_A^T M^T + M\Sigma_B V^T Y V \Sigma_B^T M^T = C. \tag{3.11}$$

Since M is nonsingular, the above equation is equivalent to

$$\Sigma_A U^T X U \Sigma_A^T + \Sigma_B V^T Y V \Sigma_B^T = M^{-1} C M^{-T}. \tag{3.12}$$

Write

$$U^T X U = \begin{array}{cl} \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{12}^T & X_{22} & X_{23} \\ X_{13}^T & X_{23}^T & X_{33} \end{pmatrix} & \begin{array}{l} r \\ s \\ n-r-s \end{array} \\ \begin{array}{ccc} r & \quad s \quad & n-r-s \end{array} & \end{array},$$

$$V^T Y V = \begin{array}{cl} \begin{pmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{12}^T & Y_{22} & Y_{23} \\ Y_{13}^T & Y_{23}^T & Y_{33} \end{pmatrix} & \begin{array}{l} p+r-K \\ s \\ k-r-s \end{array} \\ \begin{array}{ccc} p+r-k & \quad s \quad & k-r-s \end{array} & \end{array} \tag{3.13}$$

Inserting (3.6) and (3.13) into (3.12), we get

$$\begin{pmatrix} X_{11} & X_{12} S_A & 0 & 0 \\ S_A X_{12}^T & S_A X_{22} S_A + S_B Y_{22} S_B & S_B Y_{23} & 0 \\ 0 & Y_{23}^T S_B & Y_{33} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{12}^T & C_{22} & C_{23} & C_{24} \\ C_{13}^T & C_{23}^T & C_{33} & C_{34} \\ C_{14}^T & C_{24}^T & C_{34}^T & C_{44} \end{pmatrix}. \tag{3.14}$$

Therefore

$$X_{11} = C_{11}, \qquad X_{12} = C_{12} S_A^{-1}, \qquad Y_{23} = S_B^{-1} C_{23}, \qquad Y_{33} = C_{33}, \tag{3.15}$$

$$C_{13} = C_{14} = C_{24} = C_{34} = C_{44} = 0, \tag{3.16}$$

$$S_A X_{22} S_A + S_B Y_{22} S_B = C_{22}. \tag{3.17}$$

Thus, the necessary and sufficient condition (3.7) under which $\mathscr{L}_{\mathrm{II}}$ is nonempty and the expression for $\mathscr{L}_{\mathrm{II}}$ are obtained.

In addition, $\mathscr{L}_{\mathrm{II}}$ is a closed convex set, so there is a unique $[\hat{X}, \hat{Y}] \in \mathscr{L}_{\mathrm{II}}$ that makes (1.6) hold, while

$$\begin{aligned}\left\|[\hat{X}, \hat{Y}]\right\|_F^2 = {} & \|C_{11}\|_F^2 + \|C_{33}\|_F^2 + 2\|C_{12}S_A^{-1}\|_F^2 + 2\|\hat{X}_{13}\|_F^2 + 2\|S_B^{-1}C_{23}\|_F^2 \\ & + 2\|\hat{X}_{23}\|_F^2 + \|\hat{X}_{33}\|_F^2 + \|\hat{Y}_{11}\|_F^2 + 2\|\hat{Y}_{12}\|_F^2 + 2\|\hat{Y}_{13}\|_F^2 \\ & + \|S_A^{-1}C_{22}S_A^{-1} - S_A^{-1}S_B\hat{Y}_{22}S_BS_A^{-1}\|_F^2 + \|\hat{Y}_{22}\|_F^2. \end{aligned} \tag{3.18}$$

Therefore

$$\hat{X}_{13} = \hat{X}_{23} = \hat{X}_{33} = \hat{Y}_{11} = \hat{Y}_{12} = \hat{Y}_{13} = 0, \tag{3.19}$$

$$\|S_A^{-1}C_{22}S_A^{-1} - S_A^{-1}S_B\hat{Y}_{22}S_BS_A^{-1}\|_F^2 + \|\hat{Y}_{22}\|_F^2 = \min. \tag{3.20}$$

Applying Lemma 3.1 to (3.20), we obtain

$$\hat{Y}_{22} = \Phi * (S_BC_{22}S_B), \tag{3.21}$$

where Φ is defined by (3.10).

Thus (3.9) is proved by (3.15), (3.19), and (3.21). ■

Note that M is not orthogonal, so we can't discuss the least-squares problem with the GSVD.

4. SOLVING PROBLEM III

THEOREM 4.1. *Let the generalized singular-value decomposition of the matrix pair* $[A, B]$ *given in Problem* III *be of the form* (1.10). *Then*:

(a) *The set* $\mathscr{L}_{\mathrm{III\,LS}}$ *can be expressed as*

$$\mathscr{L}_{\mathrm{III\,LS}} = \left\{ M^{-T} \begin{pmatrix} \frac{1}{2}U_1^T(C+C^T)U_1 & \frac{1}{2}U_1^T(C+C^T)U_2S_A^{-1} & X_{13} & X_{14} \\ \frac{1}{2}S_A^{-1}U_2^T(C+C^T)U_1 & \frac{1}{2}\Phi * \left(S_AU_2^T(C+C^T)U_2S_A + S_BV_2^T(D+D^T)V_2\right) & \frac{1}{2}S_B^{-1}V_2^T(C+C^T)V_3 & X_{24} \\ X_{13}^T & \frac{1}{2}V_3^T(D+D^T)V_2S_3^{-1} & \frac{1}{2}V_3^T(D+D^T)V_3 & X_{34} \\ X_{14}^T & X_{24}^T & X_{34}^T & X_{44} \end{pmatrix} M^{-1}, \right.$$

$$\left. X_{13} \in R^{r\times(k-r-s)},\ X_{14} \in R^{r\times(m-k)},\ X_{24} \in R^{s\times(m-k)},\ X_{34} \in R^{(k-r-s)\times(m-k)},\ X_{44} \in \mathrm{SR}^{(m-k)\times(m-k)} \right\}, \tag{4.1}$$

where

$$\Phi = (\varphi_{ij}) \in R^{s \times s}, \qquad \varphi_{ij} = \frac{1}{\alpha_i^2 \alpha_j^2 + \beta_i^2 \beta_j^2}. \tag{4.2}$$

(b) *The set* $\mathscr{L}_{\mathrm{III}}$ *is nonempty if and only if the following three conditions hold*:

$$C = C^T, \qquad D = D^T, \tag{4.3}$$

$$CU_3 = V_1^T D = 0, \tag{4.4}$$

$$S_A^{-1} U_2^T C U_2 S_A^{-1} = S_B^{-1} V_2^T D V_2 S_B^{-1}. \tag{4.5}$$

When the above conditions are satisfied, $\mathscr{L}_{III}$ *can be expressed as*

$$\mathscr{L}_{\mathrm{III}} = \left\{ M^{-1} \begin{pmatrix} U_1^T C U_1 & U_1^T C U_2 S_A^{-1} & X_{13} & X_{14} \\ S_A^{-1} U_2^T C U_1 & S_A^{-1} U_2^T C U_2 S_A^{-1} & S_B^{-1} V_2^T D V_3 & X_{24} \\ X_{13}^T & V_3^T D V_2 S_B^{-1} & V_3^T D V_3 & X_{34} \\ X_{14}^T & X_{24}^T & X_{34}^T & X_{44} \end{pmatrix} M^{-T}, \right.$$

$$X_{13} \in R^{r \times (k-r-s)},\ X_{14} \in R^{r \times (m-k)},\ X_{24} \in R^{s \times (m-k)},$$

$$\left. X_{34} \in R^{(k-r-s) \times (m-k)},\ X_{44} \in \mathrm{SR}^{(m-k) \times (m-k)} \right\}. \tag{4.6}$$

Proof. (a): By the GSVD (1.10) of $[A, B]$ for $X \in \mathrm{SR}^{m \times m}$, we have

$$\|[A^T X A - C, B^T X B - D]\|_F^2 = \|\Sigma_A^T M^T X M \Sigma_A - U^T C U\|_F^2 + \|\Sigma_B^T M^T X M \Sigma_B - V^T D V\|_F^2. \tag{4.7}$$

Write

$$
M^TXM = \begin{array}{cc}
\begin{pmatrix}
X_{11} & X_{12} & X_{13} & X_{14} \\
X_{12}^T & X_{22} & X_{23} & X_{24} \\
X_{13}^T & X_{23}^T & X_{33} & X_{34} \\
X_{14}^T & X_{24}^T & X_{34}^T & X_{44}
\end{pmatrix} & \begin{array}{l} r \\ s \\ k-r-s \\ m-k \end{array} \\
\begin{array}{cccc} r & s & k-r-s & m-k \end{array} &
\end{array}, \tag{4.8}
$$

$$
U^TCU = \begin{array}{cc}
\begin{pmatrix}
C_{11} & C_{12} & C_{13} \\
C_{21} & C_{22} & C_{23} \\
C_{31} & C_{32} & C_{33}
\end{pmatrix} & \begin{array}{l} r \\ s \\ n-r-s \end{array} \\
\begin{array}{ccc} r & s & n-r-s \end{array} &
\end{array},
$$

$$
V^TDV = \begin{array}{cc}
\begin{pmatrix}
D_{11} & D_{12} & D_{13} \\
D_{21} & D_{22} & D_{23} \\
D_{31} & D_{32} & D_{33}
\end{pmatrix} & \begin{array}{l} r \\ s \\ p-r-s \end{array} \\
\begin{array}{ccc} r & s & p-r-s \end{array} &
\end{array}. \tag{4.9}
$$

Substituting (4.8) and (4.9) into (4.7), we obtain

$$
\begin{aligned}
&\|[A^TXA - C, B^TXB - D]\|_F^2 \\
&= \left\| \begin{pmatrix}
X_{11} - C_{11} & X_{12}S_A - C_{12} & -C_{13} \\
S_A X_{12}^T - C_{21} & S_A X_{22} S_A - C_{22} & -C_{23} \\
-C_{31} & -C_{32} & -C_{33}
\end{pmatrix} \right\|_F^2 \\
&\quad + \left\| \begin{pmatrix}
-D_{11} & -D_{12} & -D_{13} \\
-D_{21} & S_B X_{22} S_B - D_{22} & S_B X_{23} - D_{23} \\
-D_{31} & X_{23}^T S_B - D_{32} & X_{33} - D_{33}
\end{pmatrix} \right\|_F^2 .
\end{aligned} \tag{4.10}
$$

Hence, for any $X \in \mathscr{L}_{\mathrm{III\,LS}}$, the submatrices $X_{11} \in \mathrm{SR}^{r \times r}$, $X_{12} \in R^{r \times s}$, $X_{22} \in \mathrm{SR}^{s \times s}$, $X_{23} \in R^{s \times (k-r-s)}$, and $X_{33} \in \mathrm{SR}^{(k-r-s) \times (k-r-s)}$ should satisfy

$$\|X_{11} - C_{11}\|_F = \min, \qquad \|X_{33} - D_{33}\|_F = \min, \tag{4.11}$$

$$\|X_{12} S_A - C_{21}\|_F^2 + \|S_A X_{12}^T - C_{21}\|_F^2 = \min, \tag{4.12}$$

$$\|S_B X_{23} - D_{23}\|_F^2 + \|X_{23}^T S_B - D_{32}\|_F^2 = \min,$$

$$\|S_A X_{22} S_A - C_{22}\|_F^2 + \|S_B X_{22} S_B - D_{22}\|_F^2 = \min. \tag{4.13}$$

From (4.11) and (4.12), we obtain

$$X_{11} = \tfrac{1}{2}\left(C_{11} + C_{11}^T\right), \qquad X_{33} = \tfrac{1}{2}\left(D_{33} + D_{33}^T\right), \tag{4.14}$$

$$X_{12} = \tfrac{1}{2}\left(C_{12} + C_{21}^T\right)S_A^{-1}, \qquad X_{23} = \tfrac{1}{2}S_B^{-1}\left(D_{23} + D_{32}^T\right). \tag{4.15}$$

Applying Lemma 2.1 to (4.13), we have

$$X_{22} = \tfrac{1}{2}\Phi * \left[S_A\left(C_{22} + C_{22}^T\right)S_A + S_B\left(D_{22} + D_{22}^T\right)S_B\right], \tag{4.16}$$

where Φ is defined by (4.2).

Thus, the expression (4.1) for $\mathscr{L}_{\mathrm{III\,LS}}$ is obtained from (4.14), (4.15), and (4.16).

(b): From (4.10), it is easy to deduce that $\mathscr{L}_{\mathrm{III}}$ is nonempty if and only if

$$C = C^T, \qquad D = D^T, \tag{4.17}$$

$$C_{13} = C_{23} = C_{33} = D_{11} = D_{12} = D_{13} = 0, \tag{4.18}$$

$$S_A^{-1} C_{22} S_A^{-1} = S_B^{-1} D_{22} S_B^{-1}. \tag{4.19}$$

The conditions (4.18) and (4.19) are equivalent to (4.4) and (4.5) respectively. When the conditions (4.17)–(4.19) are satisfied, the expression (4.6) for $\mathscr{L}_{\mathrm{III}}$ is easily proved. ∎

Note that M is not orthogonal, so we can't discuss the minimum-2-norm solution of the equation with the GSVD.

The authors wish to thank the referee for the helpful comments.

REFERENCES

1 K. W. E. Chu, Symmetric solutions of linear matrix equations by matrix decompositions, *Linear Algebra Appl.* 119:35–50 (1989).

2 F. J. H. Don, On the symmetric solutions of a linear matrix equation, *Linear Algebra Appl.* 93:1–7 (1987).

3 G. H. Golub, and C. F. Van Loan, *Matrix Computations*, Johns Hopkins U.P., Baltimore, 1983.

4 C. G. Khatri and S. K. Mitra, Hermitian and nonnegative definite solutions of linear matrix equations, *SIAM J. Appl. Math.* 31:579–585 (1976).

5 J. R. Magnus, L-structured matrices and linear matrix equations, *Linear and Multilinear Algebra* 14:67–88 (1983).

6 J. R. Magnus and H. Neudecker, The commutation matrix: Some properties and applications, *Ann. Statist.* 7:381–394 (1979).

7 J. R. Magnus and H. Neudecker, The elimination matrix: Some lemmas and applications, *SIAM J. Algebraic Discrete Methods* 1:422–449 (1980).

8 C. C. Paige and M. A. Saunders, Towards a generalized singular value decomposition, *SIAM J. Numer. Anal.* 18:398–405 (1981).

9 G. W. Stewart, Computing the *CS*-decomposition of a partitioned orthogonal matrix, *Numer. Math.* 40:297–306 (1982).

10 W. J. Vetter, Vector structures and solutions of linear matrix equation, *Linear Algebra App.* 10:181–188 (1975).

11 J. S. Wang and X. W. Chang, Symmetric solution of a linear matrix equation, *Nanjing Univ. J. Math. Biquarterly* 7:125–129 (1990).

Received 3 July 1991; final manuscript accepted 18 November 1991